

# CS246: Mining Massive Datasets

## Homework 3

Answer to Question 1(a)

## Answer to Question 1(b)

## Answer to Question 1(c)

## Answer to Question 2(a)

## Answer to Question 2(b)

## Answer to Question 3(a)

## Answer to Question 3(b)

## Answer to Question 3(c)



## Answer to Question 4(a)

## Answer to Question 4(b)

# Cover Sheet

**Assignment Submission** Assignments are due at 11:59pm. All students (SCPD and non-SCPD) must submit their homeworks via GradeScope (<http://www.gradescope.com>). Students can typeset or scan their homeworks. Make sure that you answer each question on a separate page. Students also need to upload their code at <http://snap.stanford.edu/submit>. Put all the code for a single question into a single file and upload it. Please do not put any code in your GradeScope submissions unless explicitly asked to do so.

**Late Day Policy** Each student will have a total of *two* free late periods. *One late period extends deadlines until the day of the next lecture.* (Homeworks are usually due on Thursdays, which means the first late period expires on the following Tuesday.) However, no assignment will be accepted more than *one* late period after its due date.

**Honor Code** Students may discuss and work on homework problems in groups. This is encouraged. However, each student must write down their solutions independently to show they understand the solution well enough in order to reconstruct it by themselves. Students should clearly mention the names of all the other students who were part of their discussion group. Using code or solutions obtained from the web is considered an honor code violation. We check all the submissions for plagiarism. We take the honor code very seriously and expect students to do the same.